

## Defining PD subtypes — a step toward personalized management?

Caroline H. Williams-Gray and Roger A. Barker

**A recent article published in *Brain* proposes a clinical method for subtyping Parkinson disease cases on an individual basis, with implications for better patient stratification for personalized medicine. The authors report biological validity in terms of imaging and cerebrospinal fluid parameters, but long-term predictive validity remains to be established.**

Refers to Fereshtehnejad, S. M., Zeighami, Y., Dagher, A. & Postuma, R. B. Clinical criteria for subtyping Parkinson's disease: biomarkers and longitudinal progression. *Brain* <http://dx.doi.org/10.1093/brain/awx118> (2017)

Parkinson disease (PD) is characterized by bradykinesia, tremor and rigidity linked to the core pathology of degeneration of dopaminergic neurons within the nigrostriatal system. However, it has long been recognized that patients with PD present and progress in different ways<sup>1</sup>. An ongoing challenge is to determine whether such distinctions really constitute disease subtypes and, if so, what is their underlying pathogenic basis? Fundamentally, this debate distils down to the question of whether PD exists in different pathogenic forms, or as one pathogenic form that runs with different kinetics in different patients. In a new paper, Seyed-Mohammad Fereshtehnejad and colleagues have added to the debate by using hierarchical cluster analysis of data collected from the Parkinson's Progression Markers Initiative (PPMI) to define PD subtypes and explore their validity<sup>2</sup>.

Cluster analysis is an unbiased approach that is driven by the data, and is not influenced by *a priori* hypotheses. This technique is not new, and it has been used in a number of studies in PD since the late 1990s<sup>3–5</sup>. Such studies have tended to agree that patients with PD can be grouped into three 'types': a young-onset group with more tremor, a more benign course and fewer non-motor complications, consistent with predominant dopaminergic pathology; a group with more malignant, rapidly progressive disease consistent with widespread pathology, who are older at diagnosis; and an intermediate group. One of the key limitations of these studies is that they have described variability only at the group level. Translation of such findings to the individual patient level will be crucial to improve the stratification of patients for personalized medicine and clinical trials. Furthermore, little effort has been made to validate such subtyping at pathological and longitudinal levels.

In their study<sup>2</sup>, Fereshtehnejad and colleagues have addressed some of these limitations, first, by using a cluster analysis of the PPMI cohort to define subtypes of disease that can be applied in individual patients and, second, by probing the neurobiological basis of these subtypes using comprehensive imaging and cerebrospinal fluid (CSF) data.

The PPMI cohort comprises 'de novo' patients with PD recruited within 2 years of diagnosis and before commencing PD therapy. Hierarchical cluster analysis of baseline demographics, clinical (motor and non-motor) and neuropsychological data, and global genetic risk scores resulted in the definition of three distinct clusters of patients, which were broadly in keeping with the previous literature: a 'mild motor-predominant' group with younger onset, low motor Unified PD Rating Scale (UPDRS) scores and minimal non-motor features; a 'diffuse malignant' group with high motor and

non-motor scores; and a third 'intermediate' group. Genetic risk scores did not contribute significantly to the clustering solution.

The added value of this study derives from translation of these cluster analysis findings into a set of rules that allows stratification of every individual patient into one of the three specified groups (which is necessary because the nature of the cluster analysis approach means that not all patients will fall neatly into the statistically defined clusters). The methodology involved determining the key features that discriminated the clusters by means of principal component analysis, reducing these features into four composite domains (motor, cognition, REM sleep behaviour disorder and dysautonomia), and creating categorical definitions for subgroups, using the 75th percentile as a cut-off to define low or high composite scores.

The authors demonstrated the validity of the three defined clinical subtypes in terms of both imaging biomarkers (MRI morphometry data showed that atrophy in a PD-specific brain network was most severe in the 'diffuse malignant' group and least severe in the 'mild motor-predominant' group) and CSF biomarkers (lowest amyloid- $\beta$  (A $\beta$ ) levels and A $\beta$ :tau ratio in the 'diffuse malignant' group). Furthermore, they demonstrated that these subtypes have longitudinal validity — albeit over a short follow-up duration (average 2.7 years) — with the 'diffuse malignant' group exhibiting the fastest global progression, the greatest cognitive decline on the Montreal Cognitive Assessment, and the most rapid decline in tracer uptake on dopamine transporter single-photon emission CT.

Major strengths of the study include the *de novo* nature of the cohort (thus avoiding the confounding effect of disease stage), and the availability of such comprehensive data, including imaging and biomarker data, allowing investigation of the biological validity of the subtypes. However, an important caveat is that the outcome of any cluster analysis depends on the variables selected for inclusion — a point that is well illustrated by the differing results produced by Erro *et al.* in a cluster analysis that used the same PPMI data<sup>6</sup>. Erro *et al.* identified apathy and hallucinations as key subtype classifiers, but these variables were not included by the current authors, as they were only measured by one item on the UPDRS. Furthermore, as Fereshtehnejad *et al.* acknowledge<sup>2</sup>, the PPMI cohort is not population-representative, but biased towards individuals with higher education, greater motivation and younger age (mean 61.1 years, compared with 72.0 years in population-representative cohorts)<sup>7</sup>. It is unclear whether the under-represented patients would fit into the three-subtype solution, or whether they represent another distinct subtype (possibly an older group with relatively mild, benign disease, which is anecdotally recognizable in clinical practice).

The new study also failed to find evidence for genetic heterogeneity in association with the observed clinical heterogeneity<sup>2</sup>, which is somewhat at odds with the existing literature. In particular, an association between *GBA* mutations and more-rapid motor disease progression is well described, as are associations between *GBA*, *MAPT* and *APOE*\* $\epsilon$ 4 and cognitive decline<sup>8</sup>. This discrepancy might be explained by the fact that the genetic risk score employed in this study was derived from 28 variants associated with disease risk, which could dilute the impact of variation at any individual loci. In addition, genetic variants that underlie clinical heterogeneity might differ from those that underlie disease risk.

Despite these limitations, this study has added weight to the concept of PD subtypes. It has considerable value in providing a practical method (available as an Excel calculator) for subtyping of patients on an individual basis, and in demonstrating that these subtypes differ biologically and longitudinally. The predictive validity of this new framework now needs to be tested over longer follow-up periods, ideally in independent population-representative cohorts. Recently, the first predictive models for use in individual patients, based on a few simple baseline parameters, were

developed. Velsboer *et al.* developed an algorithm predicting 5-year outcome (dementia, postural instability or death) on the basis of age, UPDRS axial score and semantic fluency<sup>9</sup>, whereas Scherzer *et al.* developed an algorithm predicting 10-year dementia risk using *GBA* genotype, age, Mini-Mental State Examination score, education, UPDRS motor score, gender, and depression<sup>10</sup>. Both models were shown to have an accuracy of 85%.

The next challenge will be to determine the minimal data set needed to facilitate disease subtyping with the necessary fidelity for translation to the clinic, with the aim of achieving more-personalized treatments and prognoses, as well as better-stratified clinical trials.

Caroline H. Williams-Gray and Roger A. Barker are at the John Van Geest Centre for Brain Repair, Department of Clinical Neurosciences, University of Cambridge, Forvie Site, Robinson Way, Cambridge CB2 0PY, UK.

Correspondence to C.H.W.-G.  
chm27@cam.ac.uk

1. Foltynie, T., Brayne, C. & Barker, R. A. The heterogeneity of idiopathic Parkinson's disease. *J. Neurol.* **249**, 138–145 (2002).
2. Fereshtehnejad, S. M., Zeighami, Y., Dagher, A. & Postuma, R. B. Clinical criteria for subtyping Parkinson's disease: biomarkers and longitudinal progression. *Brain* <http://dx.doi.org/10.1093/brain/awx118> (2017).
3. Lewis, S. J. *et al.* Heterogeneity of Parkinson's disease in the early clinical stages using a data driven approach. *J. Neurol. Neurosurg. Psychiatry* **76**, 343–348 (2005).
4. van Rooden, S. M. *et al.* The identification of Parkinson's disease subtypes using cluster analysis: a systematic review. *Mov. Disord.* **25**, 969–978 (2010).
5. Graham, J. M. & Sagar, H. J. A data-driven approach to the study of heterogeneity in idiopathic Parkinson's disease: identification of three distinct subtypes. *Mov. Disord.* **14**, 10–20 (1999).
6. Erro, R. *et al.* Clinical clusters and dopaminergic dysfunction in de-novo Parkinson disease. *Parkinsonism Relat. Disord.* **28**, 137–140 (2016).
7. Foltynie, T., Brayne, C. E., Robbins, T. W. & Barker, R. A. The cognitive ability of an incident cohort of Parkinson's patients in the UK. The CamPaIGN study. *Brain* **127**, 550–560 (2004).
8. Collins, L. M. & Williams-Gray, C. H. The genetic basis of cognitive impairment and dementia in Parkinson's disease. *Front. Psychiatry* **7**, 89 (2016).
9. Velseboer, D. C. *et al.* Development and external validation of a prognostic model in newly diagnosed Parkinson disease. *Neurology* **86**, 986–993 (2016).
10. Liu, F. *et al.* Prediction of cognition in Parkinson's disease with a clinical–genetic score: longitudinal analysis of nine cohorts. *Lancet Neurol.* (in the press).

## Acknowledgements

The work of the authors is supported by the NIHR Cambridge Biomedical Research Centre and the Wellcome Trust–Medical Research Council Cambridge Stem Cell Institute.

## Competing interests statement

The authors declare no competing interests.

## **Author biographies**

**Caroline Williams-Gray (BMBCh MRCP PhD)** is Clinical Lecturer in Neurology at the University of Cambridge/Cambridge University Hospitals NHS Trust. She specialises in Movement Disorders. Her research focuses on defining the clinical and cognitive heterogeneity of Parkinson disease through longitudinal cohort studies, and investigating the role of the immune system in driving this heterogeneity.

**Roger A Barker (MBBS, MRCP, PhD, FMedSci)** is Professor of Clinical Neuroscience at the University of Cambridge and an Honorary Consultant Neurologist at Addenbrooke's Hospital. He combines clinical and more laboratory based research with the overarching aim of better defining subtypes of PD and Huntington's Disease (HD) and its biological basis allowing for more targeted trialling of experimental therapeutics. He has led trials of gene and cell therapies in patients with PD and HD; including the ongoing TRANSEURO trial. He currently co-leads a global initiative to deliver stem cell based therapies to patients with PD (GFORCE-PD).